Freedom as Satisfaction?

A Critique of Frankfurt's Hierarchical Theory of Freedom

Rostbøll, Christian Fogh

Published in:
Sats: Nordic Journal of Philosophy

Publication date:
2004

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Rostbøll, C. F. (2004). Freedom as Satisfaction? A Critique of Frankfurt's Hierarchical Theory of Freedom. *Sats: Nordic Journal of Philosophy*, *5*(1), 131-146. http://ej.lib.cbs.dk/index.php/sats/article/view/464/492

Freedom as Satisfaction?

A Critique of Frankfurt's Hierarchical Theory of Freedom[1]

Christian F. Rostbøll

Abstract

This article is a critical assessment of Harry Frankfurt's hierarchical theory of freedom. It spells out and distinguishes several different and irreconcilable conceptions of freedom present in Frankfurt's work. I argue that Frankfurt is ambiguous in his early formulation as to what conception of freedom of the will the hierarchical theory builds on, an avoidability or a satisfaction conception. This ambiguity causes problems in his later attempts to respond to the objections of wantonness of second-order desires and of infinite regress. With his more recent idea of freedom as being satisfied with harmony in one's entire volitional system, Frankfurt may solve the infinite regress objection but he does so at the cost of ending up with a description of freedom, which comes very close to being identical to his own description of the wanton. Frankfurt's account leaves open the question of whether the satisfactory harmony is caused by the inability to do otherwise, or is independent of it. To answer this question, Frankfurt's hierarchical theory needs to be complemented with a number of "autonomy variables" (Double). Satisfaction may be a necessary condition of freedom, but it is not sufficient. We also need to know how the person came to be satisfied. If being satisfied is merely something that happens to one, it fails to be an adequate description of a free person – and it also contradicts some of Frankfurt's own earlier insights.

I

Frankfurt first presents his hierarchical theory in his 1971 article, 'Freedom of the Will and the Concept of a Person'. The main pillar of this theory is the distinction between first-order desires and second-order desires. A first-order

[1] For comments and criticisms, I would like to thank Bernard Berofsky and an anonymous reviewer for this journal.

desire is the simple "A wants to X", where to X refers to an action. A second-order desire, on the other hand, is the reflexive "A wants to want to X", or in other words, the desire to hold a certain first-order desire. Second-order desires may be of two kinds: First, a second-order desire may be a simple desire to hold a certain desire – without being moved to act by it. Second, it may be a desire that a certain desire be my will. Frankfurt calls the second kind of second-order desire for a "second-order volition." And "it is having second-order volitions, and not having second-order desires generally, that [he] regard[s] as essential to being a person" (FWCP, p. 16),[2] and to being able to enjoy or lack free will. For Frankfurt, then, "the statement that a person enjoys freedom of the will means … that he is free to want what he wants to want. More precisely, it means that he is free to will what he wants to will, or to have the will he wants" (FWCP, p. 20). The question of freedom of the will does not concern the relation between what I want and what I do but rather the relation between my second-order volitions and my will. It is when my second-order volitions are not expressed in my will that my freedom of the will is frustrated.

The notion of a hierarchy of desires is meant to show that some desires are more truly the agent's own than others. Or that the second-order volitions in some sense are the self. We will later have to go into more depth regarding what second-order volitions more precisely are. But first let us look at the "wanton" who in contrast to a person cannot enjoy or lack free will. A wanton is an agent who does not have any second-order volitions. He does not care about his will; that is, he is indifferent as to which of his desires moves him to act.

To explicate the difference between a person who has second-order volitions and the wanton who does not, Frankfurt gives an example of two drug addicts.[3] One of the addicts hates his addiction. But we cannot say that he does not want or does not have the desire to take drugs; then he would not be an addict. However, the hierarchical theory makes it possible to say that he does not want to want to take the drug. This addict has conflicting first-order desires, but he also has a volition of the second order. Within this framework, then, we can speak of an unwilling addict. "[T]he unwilling addict may meaningfully make

[2] Abbreviations of papers by Frankfurt:
APMR – 'Alternate Possibilities and Moral Responsibility'.
FWCP – 'Freedom of the Will and the Concept of a Person'.
TCFA – 'Three Concepts of Free Action'.
I&W – 'Identification and Wholeheartedness'.
FP – 'The Faintest Passion'.

[3] Frankfurt mentions a third kind of addict, the willing addict, which I discuss below.

the analytically puzzling statements that the force moving him to take the drug is a force other than his own, and that it is not of his own free will but rather against his will that this force moves him to take it" (FWCP, p. 18). The other addict is a wanton. He does not care about being moved by his addictive desire. Because he makes no second-order volitions, there is no hierarchy among his desires – having and being moved by one desire is just as good as being moved by another for him. The wanton addict cannot be said to be a willing addict since he has no will at all. Regarding the wanton, then, the question of freedom of the will does not arise.

II

Frankfurt mentions three different conceptions of freedom in 'Freedom of the Will and the Concept of a Person', without being very clear on the distinction between the second and the third. I find it important to explicate the difference between these conceptions and will take them up in later discussions.

First, there is the idea of freedom of action:

> (1) *Freedom of action* "is fundamentally a matter of doing what one wants to do" (FWCP, p. 19).

This is clearly not what Frankfurt is concerned with in his hierarchical theory. "It misses entirely … the peculiar content of the quite different idea of an agent's whose *will* is free" (FWCP, p. 20).

We saw above that freedom of the will means having the will one wants. This statement, however, is ambiguous. And Frankfurt seems to draw two different conceptions of freedom from it. Following terms suggested by Zimmerman ('Hierarchical Motivation and Freedom of the Will', p. 356f), we might call these two conceptions, the avoidability conception and the want-satisfaction conception of freedom of the will.

The avoidability conception of freedom of the will involves the ability to do otherwise. In this it can also be called a power concept of freedom. Frankfurt is known for *not* holding avoidability as a necessary condition for moral responsibility and has made an influential argument for this view in his 1969 article, 'Alternate Possibilities and Moral Responsibility'. In that article, however, he is only concerned with moral responsibility and not with advancing a particular view of freedom. Now, consider this definition of freedom of the will from the 1971 article:

> A person's will is free only if he has the will he wants. This means that with regard to any of his first-order desires, he is free either to make that desire his will or to make some other first order desire his will

> instead. Whatever his will, then, the will of the person whose will is free could have been otherwise; he could have done otherwise than constitute the will as he did (FWCP, p. 24).

On this conception freedom of the will not only requires harmony between second-order volitions and first-order desires, but also that I could have chosen otherwise. Furthermore, this conception seems to imply that second-order volitions have *causal* efficacy. My second-order volitions have causal efficacy in this account in the sense of being able to determine which of my first order-desire becomes effective – in *making* it my will – but not in the sense of creating or generating (new) first-order desires. For short,

> (2) *Freedom of the will*, *the avoidability conception*: my will is free when and only when I have the will I want if I could have constituted it otherwise than I did.

To explain the want-satisfaction conception of freedom, consider a third kind of addict – apart from the unwilling addict and the wanton addict – the willing addict. "The willing addict's *will is not free*, for his desire to take the drug will be effective regardless of whether or not he wants this desire to constitute his will. But when he takes the drug, he takes it *freely* and *of his own free will*" (FWCP, p. 24f, emphases added). It is not entirely clear how to understand this. On the one hand, the willing addict's will is not free; on the other hand, he does have the will he wants to have. But what does it mean for a person to do something "of his own free will" when his "will is not free"? In contrast to the wanton, the willing addict has a second-order volition endorsing his first-order desire. And there is harmony between his second-order volition and his first-order desire. But the case of the willing addict does not satisfy (2), since he could not have constituted his will otherwise. Let us formulate this conception as follows:

> (3) *Freedom (of the will?), the want-satisfaction conception*: a person is free when and only when he does what he does of his own free will, even if he could not have had another will.

Now, Frankfurt is not at all clear on whether (2) or (3) is the most important for his hierarchical theory. Is it that the effective first-order desire is *caused by* – or more precisely, caused to be effective by – my second-order volition (2)? Or is it that my will is *in accordance with* my second-order volition? At one point Frankfurt seems to say that it could be either of the two: "it is in the discrepancy between his will and his second-order volitions, *or* in his awareness that their coincidence is not his own doing but only a happy chance, that the

person who does not have this freedom feels its lack" (FWCP, p. 20–1, emphasis added). But clearly it matters whether it is the one or the other.

One thing is clear, however; namely that Frankfurt thinks the weaker (3) is sufficient for attributing moral responsibility. Frankfurt sees the willing addict as responsible for taking drugs. Because the willing addict wants the desire to take drugs to be effective, "he has made this will his own. Given that it is therefore not only because of his addiction that his desire for the drug is effective, he may be morally responsible for taking the drug" (FWCP, p. 25). Frankfurt makes a distinction between what is required for freedom of the will and for moral responsibility. But moral responsibility is only of minor concern for Frankfurt in 'Freedom of the Will'; his main concern is freedom of the will. And Frankfurt only calls (2) freedom of the will.

I have dealt at some length with these distinctions because they seem to have been neglected in much of the literature,[4] and especially in order to make it clear that Frankfurt in his later elaboration and refinement of his theory seems to drop (2) in favor of (3). What I want to emphasize here is that it is (2) and not (3) Frankfurt sets forth as defining freedom of the will. And if Frankfurt drops (2), the question arises whether he is talking about freedom of the will at all.

III

We have so far neglected the question of what is so special about second-order desires and volitions. There is a reason for this: Frankfurt is in his 1971 article very vague about what makes second-order desires special. He is not clear on what makes them qualitatively different from first-order desires and thereby why they are essential for freedom of the will. There are two closely connected problems related to this vagueness. First there is the objection raised by Gary Watson: "Can't one be a wanton, so to speak, with respect to one's second-order desires and volitions?" ('Free Agency', p. 108). In other words, couldn't the second-order desires and volitions just be something one happens to have without it involving any freedom that one has them? Second, there is the problem of infinite regress. That is, the problem of how we avoid

[4] However, Kasper Lippert-Rasmussen ('Real-self Accounts of Freedom', pp. 53–58) has made a similar attempt to distinguish the different conceptions of freedom present in Frankfurt's theory. Lippert-Rasmussen conceptualizes my (2) – the avoidability conception – as "top down conformity involving alternatives" and my (3) – the want-satisfaction conception – as "mere top down conformity." I have developed my distinctions independently of Lippert-Ramussen's discussion but I do not see any contradictions between the two. Also, the way in which we use the distinctions in our overall arguments are different.

ascending to higher and higher orders to find the real self without cutting the sequence off arbitrarily.

IV

Frankfurt has tried to deal with these objections in 'Identification and Wholeheartedness' and 'The Faintest Passion'. In this section I reconstruct the argument of the first of these articles; in section V, I reconstruct the arguments of the second. I shall then, in section VI, discuss whether Frankfurt meets the objections.

Frankfurt concedes that the fact that a desire is of a higher order is not sufficient to establish its authority. He therefore takes it upon himself to show when a higher-order volition "*itself* is one by which the person *really wanted* to be determined" (I&W, p. 166). Frankfurt also notes now that our freedom is not only frustrated when there is a conflict between higher-order volitions and effective first-order desires or between volitions and will. For there is another sort of inner division. This second division manifests itself as "a lack of coherence within the realm of the person's higher-order volitions themselves" (I&W, p. 165). Whereas the first sort of division is external to the volitional complex, the second sort of division is internal to it. In the second case, it is the agent himself who is divided. The problem we face now is how we can say that a higher-order volition is truly the agent's own, that the agent is not a wanton with regard to these volitions. So for a person to be free not only requires that he acts in accordance with a higher-order volition, but also that the higher-order volition itself is one by which the person wants to be determined. The person must not only have the will he wants but also the volitions he wants.

In 'Identification and Wholeheartedness', Frankfurt puts great emphasis on the idea of making a decision. The point of decision is the point of cutting off the ascent to higher orders (I&W, p. 170). I think this idea of making a decision gives a very "activist" way of looking at freedom, and below I shall argue that it contrasts with the more passive idea of being satisfied as that which constitutes freedom.

Frankfurt wants to show that "making a decisive commitment does not consist merely in an arbitrary *refusal* to permit an interminable ascent to higher orders" (I&W, p. 167). He invokes notions like endorse, support and wholeheartedly want to designate this decisive commitment (I&W, p. 163). In order to illustrate how an agent may come to endorse or identify himself with a volition, and thereby make a decisive commitment, Frankfurt takes as an example a situation of a person who tries to solve a problem in arithmetic. The person performs a calculation, after which he may perform any number of calculations to check

the answer. If the person at some point decides for some reason to adopt a result without reservation, "he has made a genuinely unreserved commitment to the view that adopting the answer is his most reasonable alternative, he can anticipate that *this* view will be endlessly confirmed by accurate reviews of it" (I&W, p. 168). "The fact that a commitment resounds endlessly", Frankfurt continues, "*is* simply the fact that the commitment is *decisive*. For a commitment is decisive if and only if it is made without reservation, and making a commitment without reservation means that the person who makes it does so in the belief that no further accurate inquiry would require him to make up his mind" (I&W, p. 168–9). To identify with a second-order volition, then, is to commit oneself to it endlessly and without reservation.

Frankfurt argues that there is nothing arbitrary in terminating a sequence of evaluation at the point "at which there is no conflict and doubt" (I&W, p. 169). This point is the point of decision, the point of cutting off (the etymological meaning of "to decide" is "to cut off", I&W, p. 170). I said above that the solution to the problem of incoherence within a person's volitional complex requires that we can distinguish between desires that are integral to the person and those that are not. This question concerns when we can say that a desire really is the agent's own, or in other words when we can say that it is the agent who freely acts. It is at this point we can see Frankfurt's "activist" conception of freedom:

> The decision determines what the person really wants by making a desire on which he decides fully his own. To this extent the person, in making a decision with which he identifies, *constitutes himself*. The pertinent desire is no longer in any way external to him. It is not a desire he 'has' merely as a subject in whose history it happens to occur, as a person may 'have' an involuntary spasm that happens to occur in the history of his body. It comes to be a desire that is incorporated into him by virtue of the fact that he has it *by his own will* (I&W, p. 170).

Here the person constitutes himself by making a decision about his first-order desires. He is free because he has endorsed a first-order desire by an act of will as truly his. Let us try to formulate this as a definition of freedom.

> (4) *Freedom as constitutive decision*: A person is free when and only when he constitutes himself by the act of making a decision to make a desire fully his own.

This formulation should not be misunderstood so as the person can constitute himself any and all ways he likes. Frankfurt does not think we can form our character as we wish. A person is not free to causally create his own character;

he can only *take responsibility* for his characteristics by making a decisive identification with them (I&W, p. 171).

The weakness of the account focusing on constitutive decision is that it ignores the possibility of unwitting decisions with which the person does not identify.[5] Or we might say that it is possible to make decisions with which one is not satisfied. Seeing this objection, Frankfurt later abandons the notion of identification as a kind of decision in favor of an account based on the idea of being satisfied.

V

In 'The Faintest Passion', Frankfurt elaborates on the issue of volitional unity and division. Here too his concern is to avoid the problem of an arbitrary cut-off and of infinite regress but he now takes a different approach than the one focusing on the importance of making a decision. To this end he invokes the notions of wholeheartedness and ambivalence. I think we can see these two notions as analogous to the person and the wanton in 'Freedom of the Will' but at the level of higher-order, reflective attitudes. A person is wholehearted when there are no conflicts among his higher-order volitions, when he is not ambivalent. Ambivalence means that the will is divided. 'This volitional division keeps [the ambivalent agent] from settling upon or from tolerating any coherent affective or motivational identity' (FP, p. 99). About an ambivalent person there can be no truth, for there is no psychic position with which he identifies.

Apart from being the opposite of ambivalence, Frankfurt invokes the idea of being satisfied to explain what wholeheartedness means. The idea of wholeheartedness is also meant to give a fuller understanding of the notion of identification. A person is wholehearted with respect to his psychic elements when he is "fully satisfied that they, rather than others that inherently (i.e. non-contingently) conflict with them, should be among the causes and considerations that determine his cognitive, affective, attitudinal, and behavioral process" (FP, p. 103). It is important to understand exactly what being satisfied means here in order to see how Frankfurt thinks it could solve the problem concerning infinite regress. The state of being satisfied must be a state from which the person will not ascent to higher orders. Frankfurt, therefore, holds that "[b]eing genuinely satisfied is not a matter … of choosing to leave things as they are or of making

[5] For this point see Michael E. Bratman ('Identification, Decision, and Treating as a Reason', p. 193f).

some judgment or decision concerning the desirability of change. It is a matter of simply *having no interest* in making changes" (FP, pp. 104–5).

Frankfurt connects this argument to the hierarchical account in the following way. The hierarchical account builds on the idea that freedom is related to the identification with a higher-order desire endorsing a first-order desire. But the preceding argument shows, "[t]he endorsing higher-order desire must be, in addition, a desire with which the person is *satisfied*" (FP, p. 105). The conception of freedom that follows from this is something like this:

> (5) *Freedom as self-satisfaction*: I am free when and only when I am satisfied with my endorsing higher-order desires.

VI

Does Frankfurt meet Watson's charge that one can be a wanton with respect to one's second-order volitions with his idea of being satisfied? I think not. I shall argue why. In doing so I shall both go into a more detailed analysis of Frankfurt's argument in 'The Faintest Passion' and relate this argument to his earlier arguments, which I reconstructed above.

It is instructive to compare what it is to be a wanton to what it is to be a satisfied person; the one being the paradigm of non-freedom for Frankfurt, the other of freedom.[6] A wanton is an agent who has no second-order volitions, "he does not care about his will" (FWCP, p. 16). The wanton is characterized by passivity with regard to his desires. Only the person is active since it is impossible to be a passive bystander to one's second-order volitions (TCFA, p. 54). The satisfied person is satisfied with her psychic elements, which must mean that she is satisfied with the desires she has whatever order they may be. It is important for Frankfurt to construct the satisfied person so he need not form any higher-order desires in order to determine whether he is satisfied; for only in this way can he avoid infinite regress. "There is nothing that he needs to think, or to adopt, or to accept, it is not necessary for him to do anything at all" (FP, p. 104). "Perhaps his condition could be improved at no net cost, and perhaps he is aware of this, but he simply does not care" (Ibid.). This description of the satisfied person seems to come very close to the passivity characteristic of the wanton.

[6] Recall that the wanton is not even a person and, hence, cannot be unfree. To make this clear I write "non-freedom".

Frankfurt explicitly rejects this – of course. What self-satisfaction requires, according to Frankfurt,

> is that psychic elements of certain kinds *do not occur*. But while the absence of such elements does not require either deliberate action or deliberate restraint, their absence must nonetheless be reflective. In other words, the fact that the person is not moved to change things must derive from his understanding of and evaluation of how things are with him. Thus, the essential non-occurrence is neither deliberately contrived nor wantonly unselfconscious (FP, p. 105).

But what does "reflective" mean here? And how can one be reflective without recourse to higher orders? In 'Freedom of the Will', Frankfurt saw "reflective self-evaluation" as "manifested in the formation of second-order desires" (FWCP, p. 12). And in 'Identification and Wholeheartedness', he says, "it is not clear to me that adequate provision can be made for reflexivity without resorting to the notion of hierarchical ordering" (I&W, p. 165, n7). Now with regard to being satisfied he also speaks of reflection and evaluation of the self, but now it is meant to show that there is no need for recourse to higher orders. But it *is not* clear how adequate provision can be made for reflexivity without recourse to higher orders. Furthermore, how can a person evaluate "how things are with him" without "making some judgment or decision concerning the desirability of change" (FP, p. 104f)? Doesn't any evaluation involve a comparison?

Frankfurt seems to be making two points with each their purpose. These two points can be seen as two different explications of self-satisfaction in (5).

> (5a) Satisfaction with one's self does not require the agent to do anything at all.

This is meant to show that there is no danger of regress.

> (5b) Satisfaction with one's self requires the reflective absence of certain psychic elements.

This is meant to show that the satisfied person is not a wanton.

The problems I have pointed out above seem to stem from the irreconcilability of these two points.

So let us analyze closer what it is to be self-satisfied. "Satisfaction with one's self requires," Frankfurt writes, "no adoption of any cognitive, attitudinal, affective, or intentional stance … Satisfaction is a state of the entire psychic system" (FP, p. 104). That I am satisfied with my psychic state, then, cannot mean that I have formed an evaluation of this state representing my true self, nor can it mean that I desire this state. I have formed no intention about this

state, nor do I have any attitude towards it. I just happen to find myself satisfied. How can this passivity be constitutive of free will? It clearly is at odds with the activist understanding of freedom expressed in (4). On that understanding the person constitutes himself by making a decision about his first-order desires. He is free because he by an act of will has endorsed a first-order desire as truly his. But on the satisfaction account there can be no such self-constituting acts with regard to higher-order desires, for towards these there are no taking a stance. Satisfaction on Frankfurt's account is just something that happens to occur. So what does satisfaction have to do with free will? Frankfurt might object by reminding us of the requirement that absence of certain psychic elements must be reflective and self-evaluative. But I simply cannot see how this requirement is possible without taking a stance.

Now we must consider how the idea of being satisfied relates to the basic framework of the hierarchical theory, which I summarized in section I, and specifically to the different conceptions of freedom distinguished in section II above. I think this move is justified because Frankfurt does not present the idea of being satisfied as an alternative to the hierarchical account. Being satisfied, as we saw, is something the free person must be in addition to forming endorsing higher-order desires. Furthermore, there is a need to understand how the later account relates to the ambiguity in the earlier understanding.

It seems clear that the idea of being satisfied cannot be reconciled with what I called the avoidability conception of freedom of the will, (2) above. (2) is a power concept of freedom. On this view the ability to have the will I want requires that I have the power to make my first-order desires conform to my second-order volitions. This power is something I either have or do not have, if I do I am free, if I don't I am not. But now Frankfurt tells us that the higher-order desire in addition must be a desire with which the person is satisfied. If I am satisfied I am free, if I am not satisfied I am not free. But I have no power over whether I am satisfied or not. The problem is that this powerlessness goes all the way down, as it were. For the whole purpose of the idea of satisfaction was to show that my higher-order endorsements were really free, that they represented what I really wanted. Since it is only when I am satisfied with my endorsing higher-order desires that I am truly free, it is of little consequence that I have the power to have the will I want. So if satisfaction is outside my power, the consequence of Frankfurt's argument is that I have no free will. For recall that (2) was the only conception of freedom, Frankfurt calls freedom of the will.

The other candidate for a conception of freedom on the earlier account is the want-satisfaction conception, (3) above. According to (3), a person can be said

to do something freely even if he could not have done otherwise. This view seems more readily reconcilable with (5). The powerlessness of satisfaction sits more easily with the powerlessness of the willing addict. But is this a plausible understanding of freedom? I have some difficulty in seeing the willing addict as free. It seems to me that an account of freedom ought to be capable of showing that something like addiction is an obstacle to freedom. Frankfurt could show this with (2), but he has given up on this conception.

In a review of the hierarchical theory, Gary Watson points to one of its possible merits that it can show that "addictive behaviour is hardly an instance of autonomy" ('Free Action and Free Will', p. 147). According to Watson there is "a general connection" in Frankfurt's account of freedom between (i) dependency between first-order desires and second-order volitions, and (ii) alternative possibilities defined relative to higher-order volitions. But he notes that in cases of overdetermination, like in the case of the willing addict, "the connection will not hold" ('Free Action and Free Will', p. 148). I can see the appeal of the hierarchical theory with regard to the unwilling addict; in that case the hierarchical account can show how a person's freedom is frustrated by dissonance between second-order volition and effective first-order desire. But with regard to the willing addict it is hard to see how the hierarchical account could explain addictive behavior as unfree when (2) is abandoned.

VII

The uneasiness of accepting the willing addict as free can be explained in terms of a criticism of Frankfurt made by Richard Double. Double's criticism is aimed at the decisive commitment account, but I think it applies to the satisfaction account as well. Double's

> objection to Frankfurt's way of addressing the identification problem is that it at best provides only a subjective criterion for determining when choices 'belong' to agents … Frankfurtian decisive identification may be sufficient to establish that choices are *psychologically* free, that is, that agents *feel* free in making them. But decisive identification seems to go no distance toward establishing that agents are really or *normatively* free, since one can easily imagine such decisiveness exemplified by grossly irrational, unfree agents. (Double, 'A Compatibilist Account of Free will', p. 34)

For this reason Double thinks hierarchical accounts require "a strong rationality component" (Ibid. p. 35). Double develops a theory to deal with local and global controllers but I shall stick to the case of addiction. Double identifies

five autonomy variables failure among which "raises the possibility that one's choices are not under one's reasoned control" (Ibid., p. 49). These are self-knowledge, reasonability, intelligence, efficacy, and unity of agent (see summary, ibid., pp. 38, 48). I shall attempt to show what these might mean for the case of the willing addict. *Self-knowledge* requires that we know our mental states, that we have true beliefs about them. The interesting case here concerns self-deception (Ibid., p. 40). Suppose an addict believes that the only reason why he takes drugs is that he really wants to do so. This seems to be a very probable case. He believes that he has endorsed his first-order desire to take drugs freely. But in fact he has adapted his second-order volition to his irresistible first-order desire.[7] "*Reasonability* is the motivation to critically evaluate one's beliefs, desires, and choices" (Ibid., p. 38). Critically, reasonability implies that I evaluate my beliefs impartially. Thus, the addict to be reasonable must evaluate his addiction irrespectively of his inclination towards taking drugs. If his higher-order endorsement of his desire to take drugs is itself a result of his addiction he is unreasonable. "*Intelligence* is skill at gaining, retaining, and using knowledge, both about oneself and the world" (Ibid., p. 38). Regarding the addict suppose that he (1) has a first-order desire to take drugs; (2) knows that taking drugs causes early death; (3) does not want to die prematurely; (4) holds (3) stronger than (1); and (5) concludes that he endorses his addiction.[8] If one lacks the intelligence to see the contradiction in this line of reasoning, it is difficult to see what makes his endorsement a free one. "*Efficacy* is the power to control our mental states (including our choices)" (Ibid., p. 38). It "is the power to actually make self-knowledgeable, reasonable, and intelligent choices" (Ibid., p. 44). Finally "*unity* implies that there is a single agent underlying free choices" (Ibid., p. 38). It is "the requirement that the other four variables apply to the same agent" (Ibid., p. 46). I do not think this last point adds anything to Frankfurt.

Does this account give us a possibility of giving a better account of the willing addict than Frankfurt's account did? I think it does. First of all I should note that

[7] Jon Elster has written about this kind of adaptive preference formation in *Sour Grapes.* According to Elster the phenomenon of sour grapes is when an agent tries to reduce cognitive dissonance by non-consciously adapting his preferences to the possibilities (*Sour Grapes*, pp. 25, 110). Elster, however, distinguishes addiction from sour grapes. "Addiction … is much more specific than sour grapes: it is to be explained more by the nature of the object of addiction than by the tendency of the human mind to adapt whatever objects are available", (Ibid., p. 121).

[8] Double ('A Compatibilist Account of Free Will', p. 44) uses a similar example, though he forgets step (4).

the autonomy variables do not tell us whether an addict is free without a further analysis of what addiction is. If addiction in itself prevents the satisfaction of the autonomy variables, addiction clearly makes addicts unfree. Double's account helps us to give a better account of the willing addict, I think, because it supplies some questions we would like to ask about the addict. Does the willing addict deceive himself about his desires? Is his higher-order volition merely an adaptation to an irresistible first-order desire? Does he know what he is doing? Has he made an intelligent decision to take drugs? It is only when we know the answers to these question that we can make a judgment as to whether the willing addict is free or not. The uneasiness we might have about Frankfurt's bare statement that the willing addict takes drugs freely since he wants to want to take drugs (and is satisfied with this state) is a result of these questions remaining unanswered.

Double's autonomy variables may seem to be mainly epistemic in nature while I criticized the idea of being satisfied as being a passive state. But the epistemic goes together with the active in Double's account. In the case of self-deception, the agent adapts his second-order volitions unconsciously and hence passively to his first-order desire. There is no active involvement on the part of the agent. Lack of reasonability is also a state of passivity because it involves *not* critically evaluating one's desires. And so forth. It is only if we have a unified agent that reasonably, intelligently and with self-knowledge has the power to control her mental states that we can speak of free will. The problem with Frankfurt's satisfaction account is that it excludes a consideration of these active elements in identifying one's will.

VII

I have in the last few sections argued that it is difficult to see the willing addict as free just because there is no conflict between his second-order volition and first-order desire. I have insinuated that Frankfurt could have solved the problem if he had held on to (2). But I think it is also possible to make a different argument that does not rely on the possibility of having formed one's will differently.[9] We can formulate the point using Frankfurt's own point in his critique of the principle of alternative possibilities. In that connection Frankfurt makes a distinction between doing something only *because* one could not have done otherwise, and doing something one is unable to avoid doing (APMR; TCFA,

[9] As far as I can tell Double's additions do not require avoidability, but I cannot go into that here.

p. 51f). In the latter case a person does "something in circumstances that leave him no alternative to doing it, without these circumstances actually moving him or leading him to do it" (APMR, p. 2). My point is that the hierarchical account is not sufficient to making this distinction, it cannot tell when a person does something *because* of the circumstances and when he does something unavoidable independently of the fact of its unavoidability. Frankfurt, of course, thinks that when there is conflict between second-order volition and effective desire, then one does what one does because of circumstances, the person is moved to do it despite of himself. But that is not sufficient for assessing the case where there is harmony between first-order desire and second-order volition, for there is no way to tell whether this harmony is a *result* of not being able to do otherwise or is formed freely. The application of Double's autonomy variables in the preceding section showed this. It is only by investigating whether these conditions are met that we can know whether the person is passively responding to the unavoidable or actively endorsing it.

Department of Political Science
Columbia University
cfr4@columbia.edu

References

Bratman, Michael E. 'Identification, Decision, and Treating as a Reason'. In his *Faces of Intention: Selected Essays on Intention and Agency*. Cambridge: Cambridge University Press, 1999, pp. 185–206.

Double, Richard. 'A Compatibilist Account of Free Will'. In *The Non-reality of Free Will*. New York: Oxford University Press, 1991.

Elster, Jon. *Sour Grapes: Studies in the subversion of rationality*. Cambridge: Cambridge University Press, 1983.

Frankfurt, Harry. *The Importance of What We Care About: Philosophical Essays*. Cambridge: Cambridge University Press, 1988.

Frankfurt, Harry. 'Alternate Possibilities and Moral Responsibility'. In *The Importance of What We Care About: Philosophical Essays*. Cambridge: Cambridge University Press, 1988, pp. 1–10.

Frankfurt, Harry. 'Freedom of the Will and the Concept of a Person'. In *The Importance of What We Care About: Philosophical Essays*. Cambridge: Cambridge University Press, 1988, pp. 11–25.

Frankfurt, Harry. 'Three Concepts of Free Action', in *The Importance of What We Care About: Philosophical Essays*. Cambridge: Cambridge University

Press, 1988, pp. 47–57.
Frankfurt, Harry. 'Identification and Wholeheartedness'. In *The Importance of What We Care About: Philosophical Essays*. Cambridge: Cambridge University Press, 1988, pp. 159–176.
Frankfurt, Harry. 'The Faintest Passion'. In *Necessity, Volition, and Love*. Cambridge: Cambridge University Press, 1999, pp. 95–107.
Lippert-Rasmussen, Kasper. 'Real-self Accounts of Freedom'. *Sats – Nordic Journal of Philosophy*, 3. No. 2 (2002): 50–72.
Watson, Gary. 'Free Agency'. In *Free Will*. Ed. Gary Watson. Oxford: Oxford University Press, 1982, pp. 96–110.
Watson, Gary. 'Free Action and Free Will'. *Mind*, 96 (1987): 145–172.
Zimmerman, David. 'Hierarchical Motivation and Freedom of the Will'. *Pacific Philosophical Quarterly*, 62 (1981): 354–368.